

EU-Labour Force Survey November 2010 release

Setup for importing the Anonymised Yearly Data Sets for 2009

Content

- I. Overview
- II. Explanatory notes on the recoding of alphanumeric variables into numeric variables
- III. Explanatory notes about the recoding of missing values
- IV. Explanatory notes about specific recoding and the definition of variable and value labels
- V. Notes on irregularities not mentioned in the User Guide
- VI. Classifications

I. Overview

The following information will help you to import the EU-LFS 2009 yearly data. The data you received from Eurostat are formatted as comma separated values (csv). With the routines provided by GESIS's German Microdata Lab, these data can be converted into SPSS.

The routine to convert the data into SPSS consists of two files:

Setup_EULFS_2009_y.sps

Labels_EULFS_2007-2009.sps

The setup provided by the German Microdata Lab achieves the following objectives:

- Transformation of EU-LFS 2009 yearly data from csv into sav (SPSS).
- Recoding of alphanumeric variables into numeric variables.
- Recoding of missing values.
- Definition of variable and value labels (by executing Labels_EULFS_2007-2009.sps).

For a first check of the converted data, you can compare the number of cases per country with the number of cases listed in table 1. Consistent numbers indicate that the execution of the setup was successful.

II. Explanatory notes on the recoding of alphanumeric variables into numeric variables

– COUNTRYW, COUNTRY1Y

The first two numbers indicate the region, according to the country codification from 2010.¹ Countries not explicitly assigned to one region by this classification have been classified by GESIS.

The last two digits display the order for the country in alphabetical order as given by the ISO 3166-1-alpha-2 code elements (e.g. Algeria=0501).² The last two digits for countries which have been surveyed remain the same as in the COUNTRY Variable (e.g. Bulgaria=0103). The alphanumerical code from the EU-LFS deviated from the ISO classification in regard to two countries: XS instead of RS for Serbia and UK instead of GB for the United Kingdom.

In Slovenia, the coding of COUNTRYW in the raw data differs from the other countries (see User Guide 2010). Thus the recoding within the Setup is different as well.

– HHNUM, QHHNUM, QUARTER, REGION, REGIONW, REGION1Y

These variables have a large amount of values or – in case of QUARTER – only one value. Thus they remain alphanumeric.

– YEARESID

The 5-year groups (11-14, 15-19 etc.) had been coded alphanumerically and had to be recoded into numeric values in line with the standard aggregation of AGE.

III. Explanatory notes about the recoding of missing values

Coding standard for missing values:

- -1 for "No answer"
- -2 for "Not applicable"
- -3 for "Not available"
- -4 for "Not specified"
- in exceptional cases -9 for "Otherwise" (only LEAVCLAS) and -101 for "No answer" and -102 for "Not applicable" (only LEAVTIME and STARTIME, see below)

In SPSS these missing values are declared as user-missing values.

The Code -3 for "Not available" is assigned, if a variable has no valid cases in a country. Some of these variables have only system-missings (i.e. blanks) in the raw data of the respective

¹http://circa.europa.eu/irc/dsis/employment/info/data/eu_lfs/LFS_MAIN/Coding_lists_explanatory_notes_and_classifications/Country_codification_from_2010_onwards.pdf

² http://www.iso.org/iso/country_codes/iso_3166_code_lists/country_names_and_code_elements.htm

country, others have only cases with "No answer" and/or "Not applicable". For detailed information on which variables have no valid cases in which countries and have thus been recoded to -3 see table 2.

Specifics and exceptions:

– COUNTRYB, NATIONAL

The code -1 includes not only cases without a valid answer, but also suppressed cases, cases with the original code "Other or Stateless", and in some countries certain other cases (see User Guide 2010, p. 67).

– HWUSTRAI, LENTR4W, PURP4W, NA11S, NA112JS, NA11PRS, NA111YS, NA111Y1D, NA111D, NA112J1D, NA11PR1D

These variables have no valid cases in the anonymised data sets of all countries. Thus they have been recoded to -3.

– LEAVTIME, STARTIME

These variables had valid negative values in previous years. Thus the missing value category is set to -101 (for "No answer") and -102 (for "Not applicable").

– MARSTAT

In France and Slovenia children under the age 15 had not been asked about their marital status. Thus -1 for "No answer" has been recoded to -2 for "Not applicable" in those cases.

IV. Explanatory notes about specific recoding and the definition of variable and value labels

– HATFIELD

The variable contains values, which are not mentioned in the User Guide. These were labelled according to the ISCED 97 Classification of Fields of Education and Training³

– HHCOMP

The variable contains the value 50, which is not mentioned in the User Guide. All persons with this value are children with no adult household members present in the dataset. Thus we have labelled the value "No adult household in the dataset".

– ISCOPR3D, ISCO3D

A number of cases belonging to the category "armed forces" had been coded incorrectly in the 3-digit ISCO code. This became obvious while crosstabulating the 1-digit with the 3-digit ISCO code. Thus these cases have been recoded.

– REGION, REGIONW, REGION1Y

The values of these variables have country-specific meanings (according to the NUTS classification)⁴. Thus they remain unlabelled.

³http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_NOM_DTL_LINEAR&StrNom=CL_FIELD99&StrLanguageCode=EN&Int

⁴ http://circa.europa.eu/irc/dsis/employment/info/data/eu_lfs/LFS_MAIN/Coding_lists_explanatory_notes_and_classifications/NUTS_2006_from_2007_onwards.pdf

V. Notes on irregularities not mentioned in the User Guide

- AGE

No persons under the age of 15 and over the age of 74 were surveyed in Iceland, Norway and Sweden.

- AGERESID, YEARESID

There are only blanks in Iceland for persons who have not been born in Iceland. Thus there are only two categories (0 "Born in the country" and -1 "No answer").

- COEFF

There are a number of cases with a weighting factor of zero in France, Poland and Spain. There are some system-missings in Finland and the UK.

- COURLEN

There is an unusual high frequency of cases with 300 hours in France.

- HHSEQNUM

In Denmark, Iceland and Sweden all cases have the value 1. There are some system-missing in Germany and Denmark . The origin of these missings is unclear.

- INTWEEK

There are some system-missings in Denmark, France, Norway, Portugal and Sweden. The origin of these missings is unclear.

- NATIONAL

There are some system-missings in France. The origin of these missings is unclear.

- NOWKREAS

There are some system-missings in Iceland. The origin of these missings is unclear.

VI. Classifications

- ISCED classification (HATLEVEL):
http://circa.europa.eu/irc/dsis/employment/info/data/eu_lfs/LFS_MAIN/Related_documents/ISCED_EN.htm
- other classifications used in the EU LFS:
http://circa.europa.eu/irc/dsis/employment/info/data/eu_lfs/LFS_MAIN/Coding_lists_explanatory_notes_and_classifications/Index_EU_LFS_coding_%20lists_explanatory_notes_and_classifications.htm

GESIS - Leibniz-Institute for the Social Sciences
German Microdata Lab
Square B2, 1
68159 Mannheim
Tel: 0621-1246-265
Fax: 0621-1246-100
<http://www.gesis.org/gml>
E-Mail: gml@gesis.org

Table 1: Number of cases per country in the EU-Labour Force Survey 2009,
yearly data^a

Country	N (without selection)
01 AT Austria	185.242
02 BE Belgium	104.395
03 BG Bulgaria	35.329
04 CY Cyprus	41.422
05 CZ Czech Republic	235.119
06 DE Germany	50.050
07 DK Denmark	118.732
08 EE Estonia	19.856
09 ES Spain	113.337
10 FI Finland	59.445
11 FR France	417.135
12 GR Greece	297.460
13 HU Hungary	289.049
14 IE Ireland	273.130
15 IS Iceland	12.368
16 IT Italy	659.561
17 LT Lithuania	66.803
18 LU Luxembourg	19.229
19 LV Latvia	40.424
20 NL Netherlands	93.491
21 NO Norway	20.422
22 PL Poland	212.420
23 PT Portugal	164.964
24 RO Romania	239.426
25 SE Sweden	200.127
26 SI Slovenia	65.919
27 SK Slovak Republic	105.014
28 UK United Kingdom	185.764

^a see also "datafileinfo_year_2009.csv" on the CD you received from Eurostat

Table 2: Variables without valid cases in several countries,^a EU LFS 2009, yearly data

	01 AT	02 BE	03 BG	04 CY	05 CZ	06 DE	07 DK	08 EE	09 ES	10 FI	11 FR	12 GR	13 HU	14 IE	15 IS	16 IT	17 LT	18 LU	19 LV	20 NL	21 NO	22 PL	23 PT	24 RO	25 SE	26 SI	27 SK	28 UK
COUNTR1Y																					x							
COURFILD					x		x			x				x	x							x		x		x		
COURPURP					x		x															x			x			
COURWORH			x		x	x	x		x													x		x				
DEGURBA															x							x						
EDUCFILD					x	x									x							x		x		x		
EDUCLEVL															x													
EVENWK																							x					
FTPTRAS																											x	
HATLFATH							x			x					x							x			x			
HATLMOTH							x			x					x							x			x			
HATYEAR																						x						
HHAGE14							x			x					x							x			x			
HHAGEYG							x			x					x							x			x			
HHCHILDR							x			x					x							x			x			
HHCOMP							x			x					x							x			x			
HHINST	x	x	x		x		x		x	x		x	x	x		x	x	x	x	x	x	x	x	x	x	x	x	x
HHLINK							x			x					x							x			x			
HHNB0014							x			x					x							x			x			
HHNBOLD							x			x					x							x			x			
HHNBCH2							x			x					x							x			x			
HHNBCH5							x			x					x							x			x			
HHNBCH8							x			x					x							x			x			
HHNBCH11							x			x					x							x			x			
HHNBCH14							x			x					x							x			x			
HHNBCH17							x			x					x							x			x			
HHNBCH24							x			x					x							x			x			
HHNBCHLD							x			x					x							x			x			
HHNBEMPL							x			x					x							x			x			
HHNBINAC							x			x					x							x			x			
HHNBERS							x			x					x							x			x			
HHNBUNEM							x			x					x							x			x			
HHNBWORK							x			x					x							x			x			
HHFATH							x			x					x							x			x			
HHMOTH							x			x					x							x			x			
HHPARENT							x			x					x							x			x			
HHPARTNR							x			x					x							x			x			
HHPERS							x			x					x							x			x			
HHSPOU							x			x					x							x			x			
HHWKSTAT							x			x					x							x			x			
HWOVERP			x																									
HWOVERPU			x																									
INCDECIL	x				x				x					x	x						x	x			x	x		
INTWAVE						x																						
ISCOPR1D											x																	
ISCOPR3D											x																	
MAINSTAT						x																x						x
NACEPR1D																						x						
NACE1Y1D														x														

	01 AT	02 BE	03 BG	04 CY	05 CZ	06 DE	07 DK	08 EE	09 ES	10 FI	11 FR	12 GR	13 HU	14 IE	15 IS	16 IT	17 LT	18 LU	19 LV	20 NL	21 NO	22 PL	23 PT	24 RO	25 SE	26 SI	27 SK	28 UK	
PRESEEK											x			x															
REGION																					x								
REGIONW																					x								
REGION1Y															x						x	x							
REGISTER	x													x															x
SIGNISAL											x ^b																		
STAPROPR						x								x															
STAPRO1Y														x															
TEMPAGCY			x												x														
TEMPREAS																										x			
WAYJFOUN																			x										
WSTAT1Y														x															

^a Variables with either only system-missings (i.e. blanks) in the raw data or only cases with "No answer" and/or "Not applicable".

^b In the data of France there were only cases in the two categories "Don't know" and "Not applicable", what indicates that the information is not available.